



PvuI (7)
Sgfl (6) 1 GGATCTGCGATCGCTCCGGTGCCCGTCAGTGGGCAGAGCGCACATCGCCACAGTCCCCGAGAAGTTGGGGGAGGGGTCGGCAATTGAACGGGTGCCTA MfeI (82) EcoNI (96)

101 GAGAAGGTGGCGCGGGTAAACTGGGAAAGTGTGCTGTACTGGCTCCGCCCTTTTCCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

HindIII (245)
Psp1406I (203) PvuII (239) **Bsu36I (291)**
201 GTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACAGCTGAAGCTTCGAGGGCTCGCATCTCTCTTCACGCGCCCGCCCTACCTGAGGGC EcoNI (287)

301 GCCATCCACGCCGGTTGAGTCGCGTTCTGCCGCTCCCGCTGTGGTGCCTCTGAAGTGCCTCGCCGCTAGGTAAGTTTAAAGTGCAGTGCAGACC

NgoMI (441)
NaeI (441)
401 GGGCCTTTGTCGGCGCTCCCTTGAGGCTACCTAGACTCAGCCGGCTCCACGCTTTGCTGACCTGCTTGTCAACTCTACGCTTTGTTTCGTT

NcoI (560)
BstEII (555)
AgeI (552)
501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGGCCCTACCTGAGATCACCGTCACCATGGGAAAAAATACAAGAACATTGTTCTACTAAAAGG
1 M G K K Y K N I V L L K G

601 ATTAGAGGTCATCAATGATTATCATTTTGAATGGTTAAGTCTTACTGAGCAACGATTTAAAACCTAATTTAAAAATGAGAGAAGAGTATGACAAAATT
13 L E V I N D Y H F R M V K S L L S N D L K L N L K M R E E Y D K I

BstBI (766)
701 CAGATTGCTGACTTGATGGAAGAAAAGTCCGAGGTGATGCTGGTTTGGGCAAATAAAAAATTTTGAAGATATACCAACGCTTGAAGACCTGGCTG
47 Q I A D L M E E K F R G D A G L G K L I K I F E D I P T L E D L A
801 AAACCTTAAAAAGAAAAGTAAAAGTAAAAGGACCGCCATCAAGAAAGAGGAAGGAAAGTGGATGCTACTTACCTGCACCTCCACAAGCAG
80 E T L K K E K L K V K G P A L S R K R K K E V D A T S P A P S T S S
901 CACTGTCAAACTGAAGGAGCAGAGGCAACTCTGGAGCTCAGAAAAGAAAAAATCAACAAAGAAAAGGCTGGACCAAAGGAGTAAAGTGTCCGAG
113 T V K T E G A E A T P G A Q K R K K S T K E K A G P K G S K V S E

PstI (1022)
SdaI (1021) **NgoMI (1030)** **NcoI (1046)**
SbfI (1021) **NaeI (1030)** **BstXI (1040)**
1001 GAACAGACTCAGCCTCCCTCTCGAGGAGCCGGCATGTCACAGCCATGGGCGCTTCCCATCTCCCAAGACCTATTGTCAGCTCCACCCAAACACTT
147 E Q T Q P P S P A G A G M S A T G M G R S P S P K T S L S A P P N T

MscI (1124)
1101 CTTCAACTGAGAACCCGAAAACAGTGGCCAAATGTCAGGTAATCCCAGAAGAAATGTTCTCAAAAACGCCAGTGATAGTGAAGGTACTGAGTACAAC
180 S S T E N P K T V A K C Q V T P R R N V L Q K R P V I V K V L S T T
1201 AAAGCCATTTGAATATGAGACCCAGAAAATGGAGAAAAAATAATGTTTCATGTACAGTGGCTACACAGACACAGTCTTCCATGTGAAGGTTTAAAC
213 K P F E Y E T P E M E K K I M F H A T V A T Q T Q F F H V K V L N
1301 ACCAGCTTGAAGGAGAAATCAATGGAAGAAAATCATCATCATATCAGATTATTTGGAATATGATAGTCTCCTAGAGGTCAATGAAGAATCTACTGTAT
247 T S L K E K F N G K K I I I I S D Y L E Y D S L L E V N E E S T V

Psp1406I (1419) **HindIII (1490)**
1401 CTGAAGCTGGTCTTAACCAACGTTTGAGGTTCCAAATAAATCATCAACAGAGCAAAGGAAACTCTGAAGATTGATATTCTTCAAAAACCAAGCTTCAGG
280 S E A G P N Q T F E V P N K I I N R A K E T L K I D I L H K Q A S G
1501 AAATATTGTATATGGGGTATTTATGCTACATAAGAAAACAGTAAATCAGAAGACCACAATCTACGAAATTCAGGATGATAGAGGAAAATGGATGTAGTG
313 N I V Y G V F M L H K K T V N Q K T T I Y E I Q D D R G K M D V V

Tth111I (1607)
1601 GGGACAGGACAATGTCACAATATCCCTGTGAAGAGGAGATAAGCTCCAACCTTTTCTGCTTTCGACTTAGAAAAAGAACAGATGTCAAACTGATT
347 G T G Q C H N I P C E E G D K L Q L F C F R L R K K N Q M S K L I

Ppu10I (1705)
NsiI (1705)
1701 CAGAAATGCATAGTTTTTCCAGATAAAGAAAAAACAACCCGAGAAAACAATGACCCCAAGAGCATGAAGTACCCAGGAACAGAGTCAGCTTCCA
380 S E M H S F I Q I K K K T N P R N N D P K S M K L P Q E Q S Q L P N
1801 TCCTTCAGAGGCCAGCACACCTTCCCTGAGAGCCATCTCGACTCCTCAGATGCCACCAACAATCCATCCAGCAGTTTCTTCCACCAAGAAAAGTGAA
413 P S E A S T T F P E S H L R T P Q M P P T T P S S S F F T K K S E

XmnI (1917) **PvuII (1991)**
1901 GACACAATCTCAAAAATGAATGACTTCATGAGGATGCGAGATACTGAAGGAAGGGAGTCATTTTCCAGGACCGTTCATGACCAGCATAGGCCAGCTGAGA
447 D T I S K M N D F M R M Q I L K E G S H F P G P F M T S I G P A E
2001 GCCATCCCCACACTCCTCAGATGCCTCATCAACCAAGCAGCAGTTTCTTAACCAGTGAACCAAGACTGAAGACTGAACCTGAAGAAGTTTCCAT
480 S H P H T P Q M P P S T P S S S F L T T L K P R L K T E P E E V S I
2101 AGAAGCAGTGGCCAGAGTGACCTCAAAGAAGTGTGGTGTGAACGCAACAGAATCATTTGTATATGAGCCAAAGAGCAGAAGAAAATGTTTCATGCC
513 E D S A Q S D L K E V M V L N A T E S F V Y E P K E Q K K M F H A
2201 ACAGTGGCAACTGAGAATGAAGTCTCCGAGTGAAGGTTTTAATATTGACCTAAAAGGAGAAGTTCACCCAAAGAAGATCATTGCCATAGCAAAATTATG
547 T V A T E N E V F R V K V F N I D L K E K F T P K K I A I A N Y
2301 TTTGCCGAATGGGTTCTGGAGGTATATCCTTTCACACTTGTGGCTGATGTGAATGCTGACCGAAACATGGAGATCCAAAAGGATTGATTAGAAGTGC
580 V C R N G F L E V Y P F T L V A D V N A D R N M E I P K G L I R S A

EcoRI (2494)
2401 CAGCGTAACTCTAAAATCAATCAGCTTTGCTCACAACATAAAGGAAGTTTGTGAATGGGGTGTGGAGGTACATAAGAAAAATGAAGGGGTGAATTC
613 S V T P K I N Q L C S Q T K G S F V N G V F E V H K K N V R G E F
2501 ACTTATTATGAAATACAAGATAATACAGGGAAGTGAAGTGGTGCATGGACGACTGACCACAATCAACTGTGAGGAAGGAGATAAACTGAAACTCA
647 T Y Y E I Q D N T G K M E V V G V H G R L T T I N C E E G D K L K L

BglIII (2644)
2601 CCTGCTTGAATTGGCACCAGAAAAGTGGGAATACCGGGGAGTTGAGATCTGTAATTCATAGTACATCAAGGTCAATCAAGACCAGGAAAAACAAGAAAGA
680 T C F E L A P K S G N T G E L R S V I H S H I K V I K T R K N K K D

BamHI (2749) **NdeI (2789)**
2701 CATACTCAATCTGATTCAAGTATGGAACCTCACCAGACTTTTTCTTGGATCCTATCCCTATGATGTGCCAGACTATGCTGGCTATCCATATGATGTT
713 I L N P D S S M E T S P D F F F G S Y P Y D V P D Y A G Y P Y D V

MscI (2852)

NheI (2846)

2801 CCTGATTATGCTGGATACCCCTTATGATGTGCCAGACTATGCCTAAAGCTAGCTGGCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAAC
747 ▶ P D Y A G Y P Y D V P D Y A •

HpaI (2984) MfeI (2995)

2901 AGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAAGTTAACAAACAACAAATT

EcoRI (3080)

3001 GCATTCATTTTATGTTTCAGGTTTCAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAACCTCTACAAATGTGGTATGGAAATCTAAAATACAGCATA

3101 GCAAAACTTTAACCTCCAATCAAGCCTCTACTGAATCCTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGT

3201 TTGCAGCCTCACCTTCTTTCATGGAGTTAAGATATAGTGTATTTCCCAAGTTTGAACCTAGCTCTTCATTCTTTATGTTTTAAATGCACTGACCTCC

SwaI (3333) EcoO109I (3394)

3301 CACATTCCTTTTTAGTAAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCC

3401 TTCATAATATCCCCAGTTAGTAGTTGGACTTAGGGAACAAAGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCTCTGG
141 ◀ • N R T

3501 TGTACTTGAGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCATTCACTCAATGAGCACAAGCAGTCAGGAGCATAGTCAGAGATGAGCTC
137 ◀ Y K L P I L E E I T T K V L K G N M E I L V F C D P A Y D S I L E

BstXI (3623)

3601 TCTGCACATGCCACAGGGGCTGACCACCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGC
104 ◀ R C M G C P S V V R I S R D V E D S Y P H R V A V I T D F D K Q G

StuI (3758)

3701 TTGCTCACAGCAGCCCAATGGCAATGGCTTCAGCAGACAGTACCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCAGTCTTGG
70 ◀ N S V A S G I A I A E A C V T V R G I Y A E I H V A S I I E G T K T

3801 TCCTGATGGCCGCCGACATGGTCTTGTCTCATAGAGCATGGTATCTTCTCAGTGGCAGCTCCACCAGCTCCAGATCTGCTGAGAGATGTT
37 ◀ R I A A G V H H K N D E Y L M T I K E T A V E V L E L D Q Q S I N

XmnI (3900) AseI (3966)

3901 GAAGTCTTCATGATGGCCCTCTATAGTGAGTCGATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCTCTCCA
4 ◀ F T K M

4001 GCTTATCTGACGGTTCACATAACGAGCTCTGCTTATATAGACCTCCACCCTACACGCTACCGCCATTTCGCTCAATGGGGCGGAGTTGTTACGACAT

SpeI (4121)

4101 TTTGAAAGTCCCCTTGATTTACTAGTCAAAACAACTCCATTGACGTCAATGGGGTGGAGACTTGAAATCCCCTGAGTCAAACCGCTATCCACGCC

SnaBI (4249)

4201 CATTGATGACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGACTGCCAAGTAGGAAAGTCCATAAGGTCATGACTGGGC

NdeI (4354)

4301 ATAATGCCAGGCGGCCATTTACCCTCATTGACGTCAATAGGGGGCTACTTGGCATATGATACACTTGATGACTGCCAAGTGGGCAGTTTACCCTAAA

4401 TACTCCACCATTGACGTCAATGAAAGTCCCTATTGGCGTACTATGGGAACATACGTATTATTGACGTCAATGGGCGGGGTCTTGGCGGTCAGC

PaeI (4540)

PstI (4533)

SdaI (4532)

SbfI (4532)

PciI (4550)

BspLU11I (4550)

4501 CAGGCGGGCCATTTACCCTAAGTTATGTAACGCTGAGGTTAATTAAGAACATGTGAGCAAAGGCCAGCAAAGGCCAGGAACCGTAAAAAGCGCCGG
4601 TTGCTGGCGTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAATAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACC

4701 AGGCGTTTTCCCCTGGAAGTCCCTCGTGCCTCTCCTGTTCCGACCTGCCGCTTACCAGTACCTGTCGCCCTTCTCCCTTCGGGAAGCGTGGCGCT

ApaLI (4864)

4801 TTCTCATAGCTCACGCTGATGATCTCAGTTCGGTGTAGTTCGCTCCAAGTGGGCTGTGTGCACGAACCCCGTTACGCCGACCGCTGCGCC

4901 TTATCCGGTAACTATCGTCTTGTAGTCAACCCGGTAAGACACGACTTATGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAG

5001 GCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAA

5101 AAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCCGCTGGTAGCGGTGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAA

PaeI (5280) SwaI (5289) NotI (5299)

5201 GAAGATCCTTTGATCTTTTCTACGGGCTGACGCTCAGTGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAACATTTAAATCAG

EagI (5300)

5301 CGGCCCAATAAAATATCTTTATTTTATTACATCTGTGTGGTTTTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAAACAAAACGAAACAA

5401 AACAAACTAGCAAAATAGGCTGTCCCAAGTGAAGTGCAGGTGCCAGAACATTTCTCTATCGAA